ICMPv6 (ND & MLD)

Ethernet Switch

ZyNOS 4.0

Support Notes

July 2011

ZyXEL

Unleash Networking Power

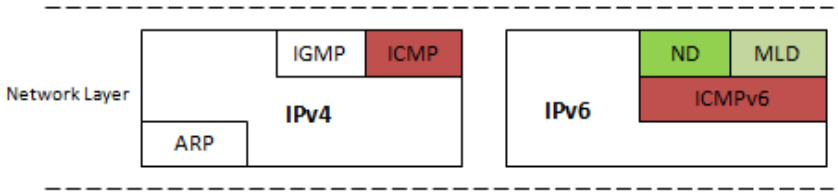
Overview of ICMPv6

Internet Control Message Protocol version 6 (**ICMPv6**) is the implementation of the Internet Control Message Protocol (ICMP) for Internet Protocol version 6 (IPv6) defined in RFC 4443. ICMPv6 is an integral part of IPv6 and performs error reporting, diagnostic functions (e.g., ping), and a framework for extensions to implement future changes.

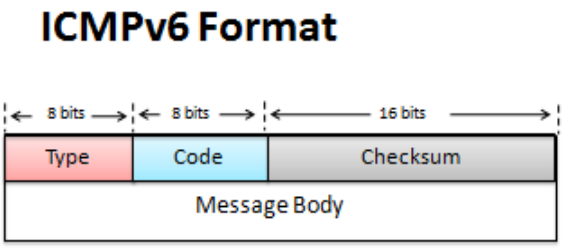
Several extensions have been published, defining new ICMPv6 message types as well as new options for existing ICMPv6 message types. Neighbor Discovery Protocol (NDP) is a node discovery protocol in IPv6 which replaces and enhances functions of ARP.Secure Neighbor Discovery Protocol (SEND) is an extension of NDP with extra security. Multicast Router Discovery (MRD) allows discovery of multicast routers.

ICMPv6 is similar to ICMPv4. It enables nodes to perform diagnostics and report problems. As in IPv4, ICMPv6 implements two kinds of messages: error messages such as "destination unreachable", "packet too big", "time exceeded" and informational messages such as "echo request" and "echo reply".

The ICMPv6 packet is identified in the Next Header field as type 58. An ICMPv6 packet is like a transport-layer packet in the sense that it is located after all the extension headers and it is the last chunk of information in the IPv6 packet.



Message General Format



Inside the ICMPv6 packet, the Type field identifies the kind of ICMP message. The Code field further details the specifics of this type of message. In order for the receiver to check the integrity of the ICMPv6 packet the Checksum field is computed over the ICMPv6 packet as well as some fields in the IPv6 header. The Data field contains information sent to the receiver for diagnostics or information purposes.

Every ICMPv6 message is preceded by an IPv6 header and zero or more IPv6 extension headers. The ICMPv6 header is identified by a Next Header value of 58 in the immediately preceding header. ICMPv6 message types:

Type		Code
1	Destination Unreachable	0 ~ 6
2	Packet Too Big	0
3	Time Exceeded	0 ~ 1
4	Parameter Problem	0 ~ 1
128	Echo Request	0
129	Echo Reply	0

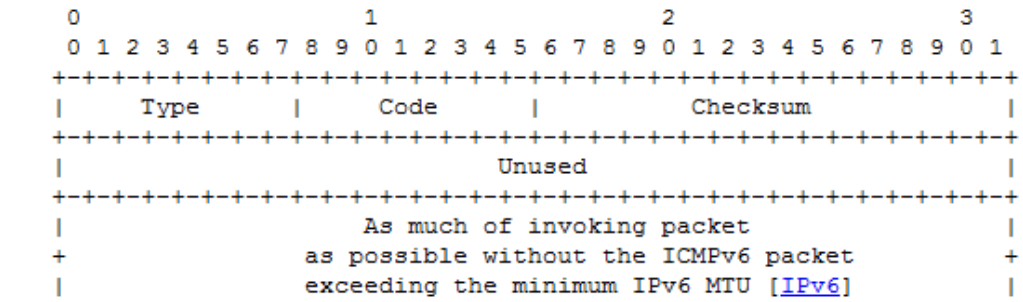
Type		Code
130	MLD (Multicast Listener Discovery) Query	0
131	MLD Report	0
132	MLD Done	0
133	Router Solicitation	0
134	Router Advertisement	0
135	Neighbor Solicitation	0
136	Neighbor Advertisement	0
137	Redirect	0
138	MLDv2 Report	0

We will implement the following ICMPv6 messages and applications in ZyXEL products. More details are described in section *錯誤! 找不到參照來源。*.

- ICMPv6 error messages
  - Destination Unreachable Message
  - Packet too Big Message
  - Time Exceeded Message
  - Parameter Problem Message
- ICMPv6 information messages
  - Echo Request Message
  - Echo Reply Message
- ICMPv6 Applications
  - Path MTU Discovery

ICMPv6 Error Messages:

Destination Unreachable Message:



IPv6 Fields:

Destination Address

Copied from the Source Address field of the invoking packet.

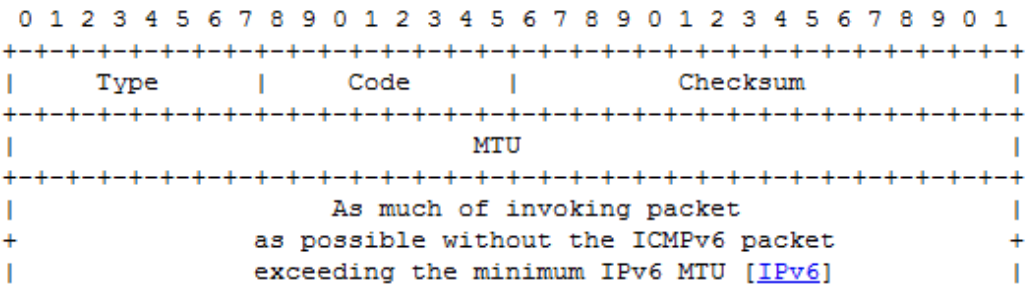
ICMPv6 Fields:

Type 1

Code
0 - No route to destination
1 - Communication with destination administratively prohibited
2 - Beyond scope of source address
3 - Address unreachable
4 - Port unreachable
5 - Source address failed ingress/egress policy
6 - Reject route to destination

Unused This field is unused for all code values. It must be initialized to zero by the originator and ignored by the receiver.

Packet Too Big Message:



IPv6 Fields:

Destination Address

Copied from the Source Address field of the invoking packet.

ICMPv6 Fields:

Type 2

Code Set to 0 (zero) by the originator and ignored by the receiver.

MTU The Maximum Transmission Unit of the next-hop link.

Time Exceeded Message:

If a router receivers a packet with a Hop Limit of zero or decrements a packet's Hop Limit to zero, it must discard the packet and originate an ICMPv6 Time Exceeded message with code 0 to the source of the packet. On the other hand, if an IPv6 node receives insufficient fragments to complete reassembly of a packet within 60 seconds of the reception of the first-arriving fragment of that packet, it must discard all the fragments and originate a Time Exceeded message with code 1 to the source of that fragment. The Time Exceeded message format is displayed in ***Figure*** *錯誤! 所指定的樣式的文字不存在文件中。**-1***.

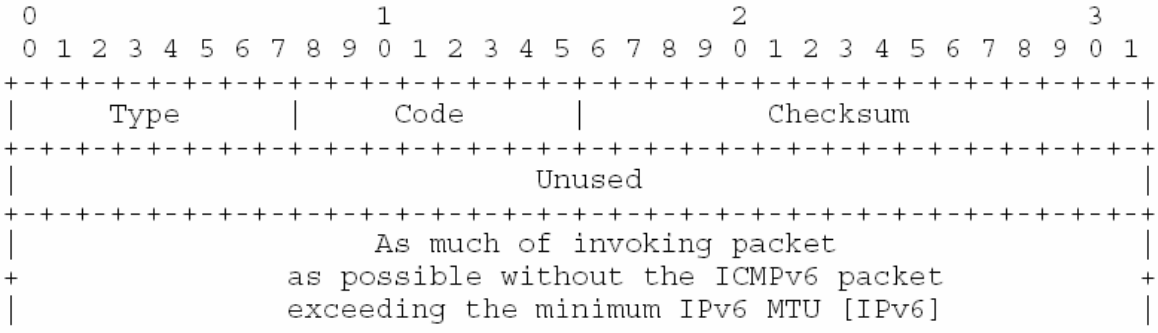


Figure *錯誤! 所指定的樣式的文字不存在文件中。**-1:** The Time Exceeded Message Format.*

For generating a Packet Too Big message, the destination address field in the IPv6

header which precedes the ICMPv6 message is first copied from the source address field of the invoking packet. The type field is set to 3. The code field is set to 0 if Hop Limit is exceeded in transit or is set to 1 for the reason of fragment reassembly timeout. The unused field is unused for all code values. It must be initialized to zero by the originator and ignored by the receiver.

Parameter Problem Message:

If an IPv6 node processing a packet finds a problem with a field in the IPv6 header or extension headers such that it cannot complete processing the packet, it must discard the packet and should originate an ICMPv6 Parameter Problem message to the packet's source, indicating the type and location of the problem. The message format is shown in ***Figure*** *錯誤! 所指定的樣式的文字不存在文件中。**-2***.

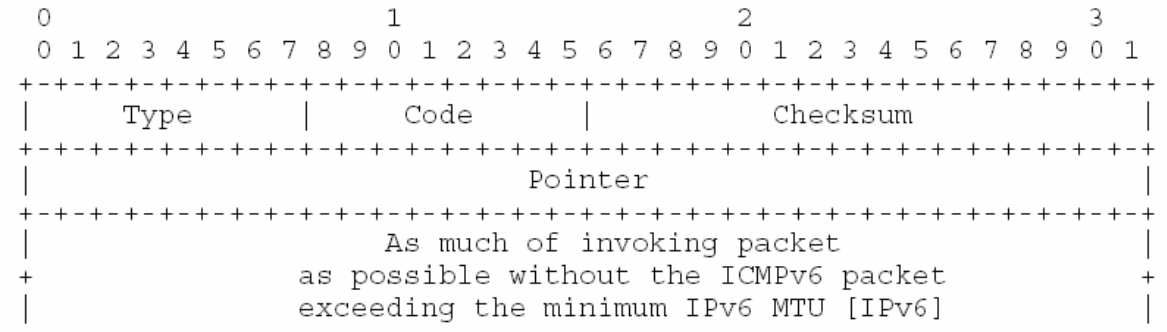


Figure *錯誤! 所指定的樣式的文字不存在文件中。**-2:** The Parameter Problem Message Format.*

The type value of a Parameter Problem message is 4. There are three kinds of code value. Code 0 indicates erroneous header field encountered. Code 1 represents unrecognized Next Header type encountered. Code 2 means unrecognized IPv6 option encountered. The pointer field is used to identify the octet offset of the original packet's header where the error was detected. The pointer will point beyond the end of the ICMPv6 packet if the field in error is beyond what can fit in the maximum size of an ICMPv6 error message.

A node receiving the above ICMPv6 error messages MUST notify the upper-layer process if the relevant process can be identified.

ICMPv6 Information Message

This section introduces the Echo Request message and Echo Reply message which comprise diagnostic functions such as ping for IPv6.

Echo Request Message

The Echo Request message format is shown in ***Figure*** *錯誤! 所指定的樣式的文字不存在文件中。**-3***.

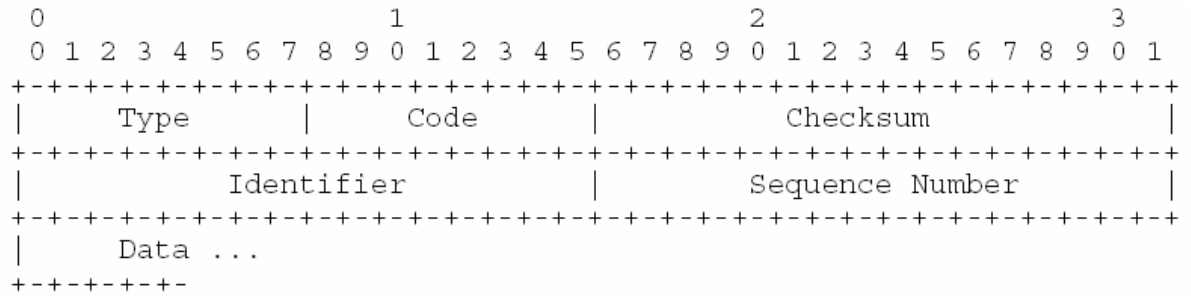


Figure *錯誤! 所指定的樣式的文字不存在文件中。**-3:** The Echo Request Message Format.*

The type of an Echo Request message is set to 128. The code value is zero. Both of the identifier field and the sequence number are to aid in matching Echo Replies to this Echo Request. Their values may be zero. The data is filled up with zero or more octets of arbitrary data. There is no limitation on the amount of data that can be put in this message.

Echo Reply Message

The Echo Reply message format is shown in ***Figure*** *錯誤! 所指定的樣式的文字不存在文件中。**-4***.

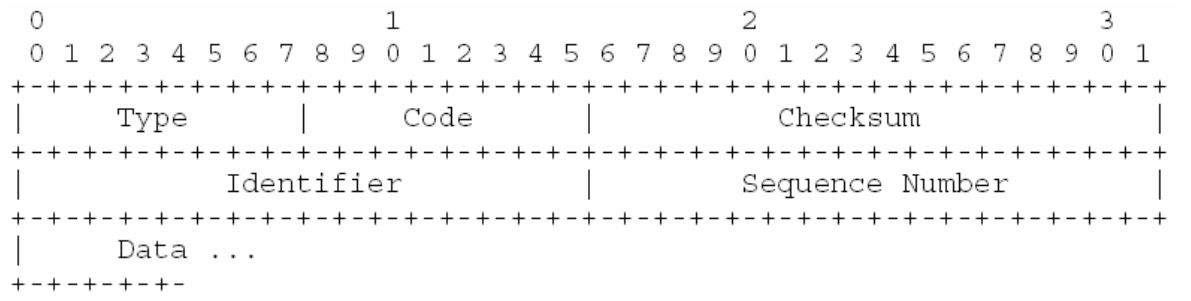


Figure *錯誤! 所指定的樣式的文字不存在文件中。**-4:** The Echo Reply Message Format.*

The destination address in the IPv6 header of this Echo Reply message is copied from the source address filed of the invoking Echo Request packet. As to the ICMPv6 message, the type is set to 129. The code value is 0. All of the identifier field, the sequence number field, and the data field are respectively copied from the invoking

Echo Request message.

It is suggested that an implementation shall perform and ICMPv6 Echo responder function that receives Echo Requests and originates corresponding Echo Replies. A node should implement an application-layer interface for originating Echo Requests and receiving Echo Replies, for diagnostic purpose. In addition, an Echo Reply should be sent in response to an Echo Request message sent to an IPv6 multicast or anycast address.

Echo Reply messages must be passed to the process that originated an Echo Request message but may be passed to those that did not originate the Echo Request message. On the other hand, Echo Request messages may be passed to processes receiving ICMP messages.

Path MTU Discovery

Protocol overview

Path MTU Discovery is a standard mechanism for a node to discover the minimum link MTU of all the links in a path between a source node and a destination node. Once a node sends packets of the largest size that will not be dropped along the path, it can conserve network resources and probably get optimal throughput.

The basic idea is that a source node initially assumes that the Path MTU of a path is the known MTU of the first hop in the path. If any of the packets sent on that path are too large to be forwarded by some node along the path, that node will discard them and return ICMPv6 Packet Too Big messages. Upon receipt of such a message, the source node reduces its assumed PMTU for the path based on the recommended MTU as reported in the Packet Too Big message.

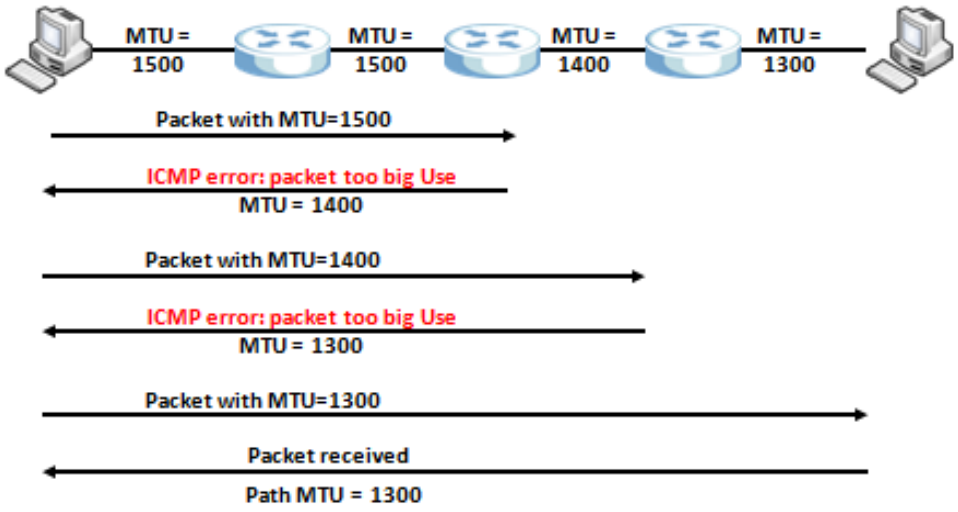


Figure 錯誤! 所指定的樣式的文字不存在文件中。-5: A Scenario of Path MTU Discovery Mechanism.

Figure 錯誤! 所指定的樣式的文字不存在文件中。-5 is a sample of this discovery mechanism. When the source node A wants to transmit a packet to the destination node D by passing through routers B and C, the first packet size A uses is 1500 bytes which is the link MTU between A and B. When B receives the packet and finds that the packet size is larger than the outgoing link MTU, B discards the packet and returns an ICMPv6 Packet Too Big error message which carries a suggested MTU for the next hop to the packet source A. Once A receives this error message, A should adjust the next packet size according to the recommended MTU in the error message, that is, 1400 bytes. After applying the same discovery steps to the router C, the source node A finally finds that the Path MTU to the destination node D is 1300 bytes. Therefore, A can transmit packets with 1300 bytes to D without being dropped.

It is convenient for a node to detect the reduction of the Path MTU by ICMPv6 Packet Too Big messages. However, there is no information for a node to increase the Path MTU. Therefore, an aging mechanism is proposed to detect increases in a path's Path MTU.

Support for Path MTU Discovery is not forced in all IPv6 nodes. It is suggested that a node which doesn't implement the mechanism send packets with the IPv6 minimum link MTU which is defined as 1280 bytes so that the packets can be forwarded to their destinations safely.

Neighbor Discovery (ND)

The Neighbor Discovery Protocol (ND) is a protocol in the Internet Protocol Suite used with Internet Protocol Version 6 (IPv6). It operates at the Network Layer of the Internet model and is responsible for address autoconfiguration of nodes, discovery of other nodes on the link, determining the Link Layer addresses of other nodes, duplicate address detection, finding available routers and Domain Name System (DNS) servers, address prefix discovery, and maintaining reachability information about the paths to other active neighbor nodes (RFC 4861).

Neighbor Discovery (ND) uses ICMPv6 messages instead of defining its own message structure. All ND messages are ICMPv6 messages types 133 (Router Solicitation), 134 (Router Advertisement), 135 (Neighbor Solicitation), 136 (Neighbor Advertisement) and 137 (Redirect). Neighbor Discovery is designed to replace ARP, ICMP Router Discovery, and ICMP Redirect used in IPv4.

Most current IPv4 implementations must be either manually configured or use a stateful address configuration protocol such as DHCP. With more computers and devices using IP, there is a need for a simpler and more automatic configuration of addresses and other configuration settings that do no rely on the administration of a DHCP infrastructure. To simplify host configuration, IPv6 supports both stateful address configuration and stateless address autoconfiguration. IPv6 stateless address autoconfiguration mechanism requires no manual configuration of hosts, minimal configuration of routers, and no additional servers.

Neighbor Discovery can commonly be divided into two parts, host and router.
ND is used by hosts to:

- Discover neighbor routers
- Discover address, address prefixes, and other configuration parameters

ND is used by routers to:

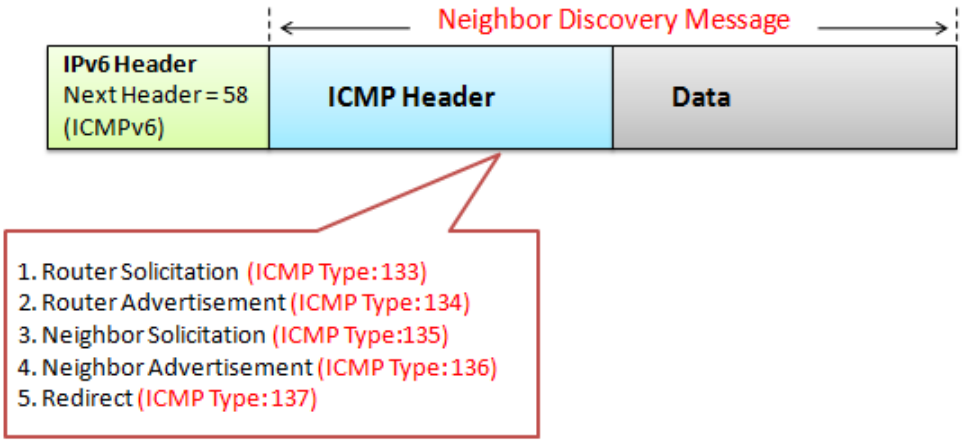
- Advertise their presence, host configuration parameters, and on-link prefixes.
- Inform hosts of a better next-hop address to forward packets for a specific destination.

ND is used by hosts and routers to:

- Resolve the link-layer address of a neighbor node.
- Determine whether a neighbor is still reachable

Regarding this specification, we only implement the host part of Neighbor Discovery.

ND Message Format



The protocol defines five different ICMPv6 packet types to perform functions for IPv6 similar to the Address Resolution Protocol (ARP) and Internet Control Message Protocol

(ICMP) Router Discovery and Router Redirect protocols for IPv4. However, it provides many improvements over its IPv4 counterparts.

Neighbor Discovery is a set of ICMPv6 messages and processes that determine relationships between neighboring nodes. Neighbor Discovery replaces ARP, ICMP Router Discovery, and ICMP Redirect used in IPv4, and Neighbor Discovery provides additional functionality.

Hosts use Neighbor Discovery to:

- Discover neighboring routers.
- Discover addresses, address prefixes, and other configuration parameters.

The following table lists and describes the Neighbor Discovery processes that RFC 2461 describes.

Neighbor Discovery Process	Description
Router discovery	The process by which hosts discover the local routers on attached links. Equivalent to ICMPv4 Router Discovery. For more information, see "Router Discovery" later in this section.
Prefix discovery	The process by which hosts discover the network prefixes for local link destinations. Similar to the ICMPv4 Address Mask Request/Reply. For more information, see "Router Discovery" later in this section.
Parameter discovery	The process by which hosts discover additional operating parameters, including the link MTU and the default hop limit for outgoing packets. For more information, see "Router Discovery" later in this section.
Address autoconfiguration	The process for configuring IP addresses for interfaces in either the presence or absence of a server that provides stateful address configuration using a protocol such as Dynamic Host Configuration Protocol version 6 (DHCPv6).
Address resolution	The process by which nodes resolve a neighbor's IPv6 address to its link-layer address. Equivalent to ARP in IPv4. For more information, see "Address Resolution" later in this section.
Next-hop determination	The process by which nodes determine the IPv6 address of the neighbor to which a packet is being forwarded based on the destination address. The forwarding or next-hop address is either the destination address or the address of an on-link default router.
Neighbor unreachability detection	The process by which nodes determine that a neighbor is no longer receiving packets.
Duplicate address detection	The process by which nodes determine whether an address considered for use is already in use by a neighboring node. Equivalent to using gratuitous ARP frames in IPv4.
Redirect function	The process of informing a host of a better first-hop IPv6 address to reach a destination. Equivalent to the use of the IPv4 ICMP Redirect message.

Multicast Listener Discovery Protocol (MLD)

The use of multicasting in IP networks is defined as a TCP/IP standard in RFC 1112, "Internet Group Management Protocol (IGMP)." This RFC defines address and host extensions for the way in which IP hosts support multicasting. The same concepts originally

developed for the current version of IP, known as IP version 4 (IPv4), also apply to IPv6.

Multicast traffic is sent to a single address but is processed by multiple hosts. Multicasting is similar to a newsletter subscription. As only subscribers receive the newsletter when it is published, only host computers that belong to the multicast group receive and process traffic sent to the group's reserved address. The set of hosts listening on a specific multicast address is called a multicast group.

Other important aspects of multicasting include the following:

- Group membership is dynamic, allowing hosts to join and leave the group at any time.
- The joining of multicast groups is performed through the sending of group membership messages. In IPv6, Multicast Listener Discovery (MLD) messages are used to determine group membership on a network segment, also known as a link or subnet.
- Groups are not limited by size and members can be spread out across multiple network segments (if connecting routers support the forwarding of multicast traffic and group membership information).
- A host can send traffic to the group's address without belonging to the corresponding group.

IPv6 multicast addressing

IPv6 multicast addresses are reserved and assigned from the Format Prefix 1111 1111 (0xFF). The following table is a partial list of IPv6 multicast addresses that are reserved for IPv6 multicasting and registered with the Internet Assigned Numbers Authority (IANA).

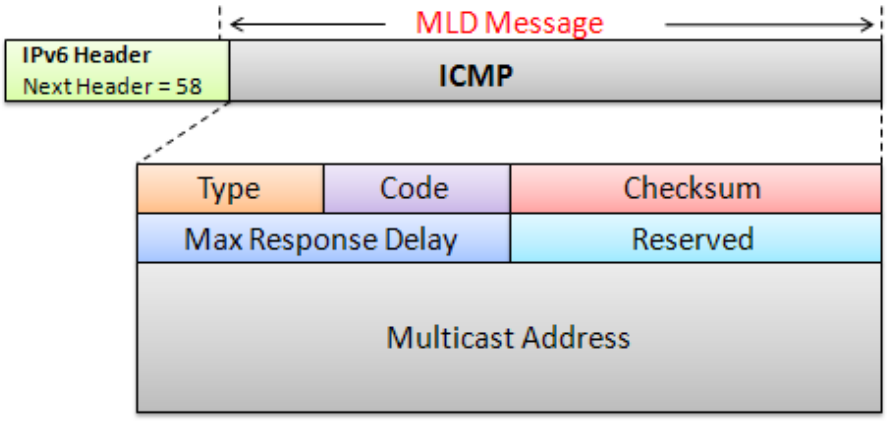
For a full and current listing of additional IPv6 addresses that are reserved for multicasting, see the Information Sciences Institute Web site document: Internet Protocol Version 6 Multicast Addresses.

IPv6 multicast address	Description
FF02::1	The all-nodes address used to reach all nodes on the same link.
FF02::2	The all-routers address used to reach all routers on the same link.
FF02::4	The all-Distance Vector Multicast Routing Protocol (DVMRP) routers address used to reach all DVMRP multicast routers on the same link.
FF02::5	The all-Open Shortest Path First (OSPF) routers address used to reach all OSPF routers on the same link.
FF02::6	The all-OSPF designated routers address used to reach all OSPF designated routers on the same link.
FF02::1:FF*XX:XXXX*	The solicited-node address used in the address resolution process to resolve the IPv6 address of a link-local node to its link-layer address. The last 24 bits (*XX:XXXX*) of the solicited-node address are the last 24 bits of an IPv6 unicast address.

A single IPv6 multicast address identifies each multicast group. Each group's reserved IPv6 address is shared by all host members of the group who listen and receive any IPv6 messages sent to the group's address.

IPv6 multicast addresses are mapped to a reserved set of media access control (MAC) multicast addresses. For information about the mapping of IPv6 multicast addresses to Ethernet MAC addresses, see RFC 2464, "Transmission of IPv6 Packets over Ethernet Networks."

MLD messages Format



Type: There are three types of MLD messages.

1) Multicast Listener Query:
   - General Query, used to learn which multicast addresses have listeners on an attached link.
   - Multicast-Address-Specific Query, used to learn if a particular multicast address has any listeners on an attached link.

2) Multicast Listener Report

3) Multicast Listener Done. When an MLDv1 host leaves a group, it sends a multicast listener done message to multicast routers on the network. The above messages types are referred to simply as "Query", "Report", and "Done".Code: Initialized to zero by the sender; ignored by receivers.Checksum: The standard ICMPv6 checksum, covering the entire MLD message plus a "pseudo-header" of IPv6 header fields.Maximum Response Delay: The Maximum Response Delay field is meaningful only in Query messages, and specifies the maximum allowed delay before sending a responding Report, in units of milliseconds. Reserved: Initialized to zero by the sender; ignored by receivers.

Multicast Address

MLD is used to exchange membership status information between IPv6 routers that support multicasting and members of multicast groups on a network segment. Host membership in a multicast group is reported by individual member hosts, and membership status is periodically polled by multicast routers. MLD is defined in RFC 2710, "Multicast Listener Discovery (MLD) for IPv6."

MLD message types are described in the following table.

MLD message type	Description
Multicast Listener Query	Sent by a multicast router to poll a network segment for group members. Queries can be general (requesting group membership for all groups), or specific (requesting group membership for a specific group).
Multicast Listener Report	Sent by a host when it joins a multicast group, or in response to an MLD Multicast Listener Query sent by a router.
Multicast Listener Done	Sent by a host when it leaves a host group and might be the last member of that group on the network segment.

MLD messages are sent as ICMPv6 messages.

Note

- IPv6 is a rapidly evolving standard. The RFCs referenced might have been made obsolete by newer RFCs.